

## PĀLI PHILOLOGY AND THE STUDY OF BUDDHISM

K.R. Norman

In the *Times* for October 1987 there appeared an article which began with these words:

“More than 5,000 manuscripts contain all or part of the New Testament in its original language. These range in date from the second century up to the invention of printing. It has been estimated that no two agree in all particulars. Inevitably, all hand written documents are liable to contain accidental errors in copying. However, in living theological works it is not surprising that deliberate changes were introduced to avoid or alter statements that the copyist found unsound. There was also a tendency for copyists to add explanatory glosses. Deliberate changes were more likely to have been introduced at an early stage before the canonical status of the New Testament was established. If one argues that no one manuscript contains the original, unaltered text in its entirety, then one cannot select any one of these manuscripts and rely exclusively on its text as if it contained the monopoly of the original words of the original authors.”

The article went on to point out that if one further argues that the original text has survived somewhere among the thousands of extant manuscripts, then one is forced to read all these manuscripts, to assemble the differences between them in a systematic way, and then to assess, variant by variant, which manuscripts have the original and which the secondary text. It is not surprising that such a prospect has daunted many biblical scholars who have been content to rely on the printed texts of earlier ages, in which the evidence of only a few favored manuscripts was used. Even many recent printed editions of the Greek New Testament, and modern translations based on these, have usually followed this practice of building their text on a narrow base that is unlikely to be entirely original. All those who read theological literature and, in particular, commentaries on books of the New Testament will be aware that the interpretation can often depend on the precise definition of a word, phrase or verse. There can be no doubt that the precise form of the original text is a matter of crucial concern.

That article was referring to the second part of an edition of the Gospel according to St Luke,<sup>1</sup> a gospel which was selected to inaugurate an enterprise intended to provide the scholarly world with a comprehensive collection of variant readings in the Greek New Testament. For that edition it was decided to display all significant variant readings in

---

<sup>1</sup> *The Gospel according to St Luke*, Part II, chapters 13-24, edited by the American and British committees of the International Greek New Testament Project, Oxford University Press, 1987.

more than two hundred of those manuscripts which contain St Luke's Gospel, as well as early translations of the Gospel, and quotations from the Gospel in the works of the early Church Fathers. At one stage more than two hundred and sixty readers were engaged in studying and collating microfilms of the Greek manuscripts utilized, and many scholars over many years have been involved in the preparation of the work.

Reading the beginning of the article I was struck by how close, *mutatis mutandis*, is the situation with regard to the books of the Pāli canon. Reading, however, about the way in which this particular edition was made, I was struck by the complete contrast to the way in which many editions of the Pāli texts have been, and are, I fear, still being, made. Leaving aside those texts which have been edited from a single manuscript because, unfortunately, only one single manuscript has so far come to light, anyone who reads the editor's preface to many of the editions published by the Pali Text Society will be amazed at the small number of manuscripts which the editors have thought would be sufficient for them to utilize when performing their task. In some cases editors have been content to reproduce the readings of one or more oriental printed editions, often without attempting to ascertain the basis for such editions. For example, the Pali Text Society edition of the *Buddhavamsa-aṭṭakathā* is based upon, and is in effect a transcription of, a single printed edition, that in Sinhalese script in the Simon Hewavitane Bequest Series. It occasionally gives variant readings from that edition. Volume I of the Pali Text Society edition of the *Papañcasūdanī*, the commentary upon the Majjhima-nikāya, is based upon two Sinhalese manuscripts, two Sinhalese printed editions and a Burmese manuscript of the ṭikā, i.e. the subcommentary upon the *Papañcasūdanī*, which could, at best, have given help with whatever words are quoted in the lemmata. From Volume II onwards the basis of the edition was three printed editions, one being one of the Sinhalese editions used for Volume I, and the other two being editions in the Burmese and Thai scripts. No information whatsoever is given about the basis for these oriental editions, nor are any variant readings quoted from them. No information is given about the principles followed in establishing the text of the Pali Text Society edition, and we are left to suppose that, when the oriental editions differed, the editor of each volume selected arbitrarily whatever readings appealed most to him or her. Other editions have been printed without the benefit of proof-reading, in part or in whole, and one was actually printed with spaces, rather than hyphens, between component parts of compounds, because the Founder's widow, acting as General Editor, was mindful of her dead husband's dislike of hyphens<sup>2</sup> and arbitrarily ordered the printer to remove all those inserted by the editor in his manuscript. This he did, but he omitted to close up the consequent gaps.

It is doubtful whether these facts are known to many of those who write about Theravāda Buddhism, and who happily base their work upon texts which have been edited in this way, and the translations based upon such texts. Even those who are aware of such deficiencies frequently do nothing about it either because they do not have the time, or because they are not sufficiently competent in the Pāli language to remedy the matter. It may justifiably be asked whether the errors which may remain in the editions of the Pāli texts really matter, and whether they are likely to have resulted in any misunderstanding of the basic and most important elements of Buddhism. My simple

---

<sup>2</sup> See M.M. Bose, ed., *Ivuttaka-aṭṭakathā*, Vol. II, London, 1936, iii.

answer is that I do not know, because I am not competent to judge the relevant importance of Buddhist doctrines, but as a matter of principle, I would regret any errors of facts, however trivial, or interpretation of those facts, if they arose from an error in an edition of a Pāli text, just as no New Testament scholar worthy of the name would be happy about anyone working with a text which he knew to be less than perfect.

It seems to me that the situation in other fields of Buddhist studies is not so very different. The main difference is that, in the area of Hīnayāna Sanskrit texts at least, the number of manuscripts concerned is much smaller, and in many cases, when we come to consider the texts from Gilgit or Turfan, we are talking about unique manuscripts or fragments of manuscripts. It is not clear that some of those making use of these manuscripts realize the implications of this. When we talk about the deficiencies of a Pāli edition based upon one or two manuscripts or printed texts, we are doing so in the knowledge that, if we compare this handful of source materials with all the manuscripts which we know to be available in the libraries of the world, such a small number is not likely to be a wide enough sample to ensure correctness. Why then should we accept that the unique Kharoṣṭhī *Dharmapada* is likely to be a correct version of the *Dharmapada* of the Dharmaguptaka school, or a section of the Mūlasarvāstivādin *Vinaya* from Gilgit represents the authentic version of that text in every respect?

We have evidence that there were variations in the versions of such texts which these schools had, as Schmithausen has shown us very recently,<sup>3</sup> and if we find such discrepancies in the few versions of any one text which the sands of Chinese Turkestan have given up, or which have come to light in Kashmir, then what would the situation be if we had a far wider and more representative sample of the literature of the Hīnayāna schools? I am well aware of the fact that scholars working in such fields sometimes say that they can compare their texts with the Tibetan and Chinese translations, and by emending them in the light of those translations they can arrive at a correct version of (say) the Mūlasarvāstivādin *Vinaya*. To them I would say that it may be possible by comparing Sanskrit, Tibetan and Chinese versions of such a text to come to an agreed reading of a particular passage, but it must be realized that in most cases the Tibetan and Chinese versions have no independent authority. They were made from Sanskrit originals, and all such a comparison can do is to confirm the reading of the Sanskrit text from which those translations were made.

In some cases it may be thought sufficient to do this, but in reality our aims should be grater than this. We know very little about the translation techniques which were adopted by those early translators and we have no idea what steps were taken to ensure that the manuscript or manuscripts from which they were making their translations contained a correct version of the text. We know from the records of the Chinese pilgrims that they sometimes obtained a single manuscript of a text to take back to China, from which in due course they or their successors made their translation. Without more information we cannot be certain that the Sanskrit (or very occasionally Pāli) version from which they made their translation was free from errors. Even if it was, then we must remember that

---

<sup>3</sup> See L. Schmithausen, "Beiträge zur Schulzugehörigkeit und Textgeschichte kanonischer und postkanonischer Materialien", in H. Bechert, ed.: *Zur Schulzugehörigkeit von Werken der Hīnayāna-Literatur*, Zweiter Teil, Göttingen, 1987, 304-81.

the Sanskrit version was in turn, a translation from some variety of Middle Indo-Aryan dialect, and even if we can establish the form of the Sanskrit version correctly, all it tells us is what the person or persons responsible for making that translation thought his Middle Indo-Aryan exemplar meant. It does not prove that he was correct in his interpretation. It cannot be emphasized too much that all the versions of canonical Hīnayāna Buddhist texts which we possess are translations, and even the earliest we possess are translations of some still earlier version, now lost.

Clearly, for the study of Theravāda Buddhism accurate editions of Pāli texts are essential. We must then face the question: “What is an accurate edition of a Pāli text?” Here the variations between regional versions may cause problems. If we find, for example, that the Burmese edition of the first verse of the *Suttanipāta* contains the word *visaṭa*, with the retroflex *-ṭ-*, while the Sinhalese edition has *visata*, with dental *-t-*, then we have to recognize the fact that we may be faced, not with a correct reading as opposed to an incorrect one, but with a fundamental dialect difference of Middle Indo-Aryan, whereby *-ṭ-* followed by a dental *-t-* may or may not change that dental *-t-* to retroflex *-ṭ-* before it disappears. Consequently both readings may be correct in Pāli, and both may be original, since both may go back to dialects of Middle Indo-Aryan which are older than Pāli, perhaps back to the time of the Buddha. In short, the Buddha may well have used both versions in different recitations of the same text in different dialects. This aspect of Middle Indo-Aryan philology has not always been clear to scholars, even very eminent scholars, and as a result we find such statements as “the alternative spelling *visaṭam*... is supported by the [Gāndhārī] Prakrit [and] should certainly be restored to the text”<sup>4</sup> with a multitude of suggestions as to how the word should be taken. In this situation we should bear in mind the fact that the redactor of the *Udānavarga*, who most likely had something very similar to the *Gāndhārī Dharmapada* as his exemplar, was able to recognize that the word was to be identified with Sanskrit *visṛta*.<sup>5</sup>

In some cases, however, the growing amount of material we have from non-Pāli sources can sometimes be used, if we exercise great care, to support one Pāli reading against another. The relationship between Pāli and non-Pāli versions of one and the same text, or phrase, or individual word, does nevertheless raise problems, since it is not at all obvious why a reading in a Sanskrit or Prakrit manuscript from Chinese Turkestan should sometimes be closer to a reading in a Pāli manuscript from Burma or Thailand than to a reading in a Sinhalese manuscript, e.g. the *Udānavarga*<sup>6</sup> has *ḥṛntana* in the verse which is parallel to *Dhammapada* 275, where the Sinhalese edition has *santhana*, but the Burmese edition has *kantana*. On the other hand, the Sanskrit version of the *Upāli-sūtra* has *aprabhītasya* where the Sinhalese and Burmese editions of the *Majjhima-nikāya*<sup>7</sup> have *appahīnassa*, but the Siamese version has *appabhītassa*.<sup>8</sup> Much research needs to be carried out into the interrelationship between the various Buddhist countries and their

<sup>4</sup> J. Brough, *The Gāndhārī Dharmapada*, Oxford, 1962, 197.

<sup>5</sup> *Udānavarga*, edited by F. Bernhard, XXXII, 64 foll.

<sup>6</sup> *Udānavarga*, *op. cit.*, XII, 9-10.

<sup>7</sup> *Majjhima-nikāya*, I 386, 25\*.

<sup>8</sup> See O. von Hinüber, “Upāli’s verses in the *Majjhimanikāya* and the *Madhyamāgama*”, in L.A. Hercus *et al.*, eds., *Indological and Buddhist Studies* (Volume in honor of Professor J.W. de Jong on his sixtieth Birthday), Canberra, 1982, 243-51 (see page 244).

manuscript traditions to try to find out the extent to which they depended upon one another in the past, in an attempt to work out how far their manuscript traditions are independent. It is clear that in very recent years the tradition in Thailand has been greatly influenced by Burmese and European editions, but research carried out in libraries in Thailand<sup>9</sup> is uncovering manuscripts which seem to be older than anything we have available from Ceylon and Burma, and some of the readings found in such manuscripts differ from those found in the present Thai editions, and give support for alternative readings which are in many ways superior to those of our present editions. These manuscripts certainly pre-date the Burmese Fifth and Sixth Councils, and in content, if not in actual physical nature, perhaps go back to the Siamese council held in 1475-77.

If non-Pāli sources can be used to help us in our research in Pāli philology, then the reverse is also the case. This has, of course, been recognized by those editing Sanskrit manuscripts from Turfan and Gilgit, and it is common practice to print the Pāli version, where it exists, alongside such a Sanskrit text. This has proved very useful as a means of correcting errors or conjecturing ways of filling up lacunae in manuscripts, or placing fragments in order, etc. The next stage of such an investigation, however, is to go further than this, and to compare the Pāli and non-Pāli versions, and to try to deduce, if not the form of the original text, at least that of an earlier version, from which they have both been translated.

Such a need arises immediately where we come across words which clearly refer to the same thing, but have different forms, which cannot be easily explained by the normal dialect variations, e.g. Sanskrit *pratisamvid*, *avadāna*, *ekavīcika*, *anupadhiśeṣa*, and *saṅghāvaśeṣa*, where the Pāli forms are *paṭṣambhidā*, *apadāna*, *ekabījīn*, *anupādisesa* and *saṅghādisesa*. If we wish to make use of etymology as a means of finding out the precise meanings of these technical terms, then the fact that the relationship between them is obscure makes our task more difficult. There are also difficulties when we come across words which are possibly ambiguous. It is well-known that certain Pāli words have two or more possible etymologies, i.e. two or more Sanskrit words have become homonymous in Middle Indo-Aryan, so that when we meet the Pāli word in our reading we have to decide which of the Sanskrit antecedents we are dealing with. It is very interesting in such contexts to find that sometimes the Sanskrit parallels do not distinguish between the alternatives, but select one or other of them, e.g. Pāli *nekkhamma* can be derived from either Sanskrit *naiṣkramya* or Sanskrit *naiṣkāmya*, but it seems always to be Sanskritised in Buddhist texts as *naiṣkramya*. Reading, therefore, a Buddhist Sanskrit text in which the word *naiṣkramya* appears, we must bear in mind that it may stand for *naiṣkāmya*.

If the original author of a text intended a pun which was possible because the two elements of his pun were homonymous in the dialect of Middle Indo-Aryan in which he was composing his text, then the redactor translating into Sanskrit was faced with a problem when he came to deal with it. If an author intended *dhamma-pīti* to mean both “drinking the doctrine” and “joy in the doctrine”, then a Sanskrit redactor, even if he realized that a pun was intended, which is not at all certain, could not hope to express it in

---

<sup>9</sup> See O. von Hinüber, “Pāli manuscripts from canonical texts from North Thailand – a preliminary report”, JSS, 71, 1983, 75-88.

Sanskrit, since he had to make a choice between writing *dharmā-pīti* and *dharmā-prīti*. He had the same problem with regard to a pun based upon *atta-dīpa*, which could mean either “a lamp for oneself” or “an island, i.e. refuge, for oneself”. He had to write either *ātma-dīpa* or *ātma-dvīpa*.

Such examples are well-known, but there are other forms, equally ambiguous, which are perhaps less well known. If we consider the Sanskrit word *bodhisattva*, I do not doubt that many people would translate it as “a being destined for enlightenment”, and the same translation is usually given for the Pāli form *bodhisatta*. Monier-Williams, however, translates it more according to the rules of Sanskrit grammar, as “one whose essence is perfect knowledge”. This is a very good epithet for the Buddha, but hardly suitable for one who has not yet reached that state, which would make us rather suspicious about the translation of the word. We can, however, point out that the word *bodhisattva* is late in Sanskrit, and probably later than *bodhisatta* in Middle Indo-Aryan. We can therefore postulate that it is a backformation in Sanskrit. This gives us the opportunity of proposing alternative etymologies, and we can, if we wish, accept the suggestion of the Pāli commentators that it is *bodhi + satta < śakta*, i.e. “capable of enlightenment”.<sup>10</sup>

Sometimes Pāli philology can help to suggest a solution to problems in languages other than Pāli. Those of you who have read Nāgārjuna’s *Ratnāvalī* in Tibetan may have noticed that he refers to *nirvāṇa* as ‘master of everywhere’<sup>11</sup>, while the Chinese version translates it as ‘all pervading.’<sup>12</sup> The epithet is a quotation of a canonical phrase, which appears in the Chinese translation of the *Dīrgha-āgama*<sup>13</sup> of the Dharmaguptakas in the form ‘shining of or by itself’, although the parallel in the Chinese translation of the *Madhyama-āgama*<sup>14</sup> of the Sarvāstivādins of Kāśmīra seems not to include the epithet. The phrase also occurs twice in the Theravādin canon, in the *Dīrgha-nikāya*<sup>15</sup> and the *Majjhima-nikāya*,<sup>16</sup> and most editions read *sabbatopabha*,<sup>17</sup> which would appear to support the reading in the *Dīrgha-āgama*. Why, then, should Nāgārjuna, or at least his translators, translate differently?

When, however, we come to investigate, we find that the Pāli situation is not as simple as it might appear. The commentator Buddhaghosa wrote commentaries on both the *Dīrgha-* and the *Majjhima-nikāya*. In the commentary on the latter he gives three explanations for *sabbato-pabha*: ‘shining’, ‘abundant, having power’, and ‘ford’. The sub-commentary to his commentary refers only to the first of these. In his commentary upon the *Dīrgha-nikāya* Buddhaghosa gives only the explanation as ‘ford’, basing it upon the sound change *-p- > -bh-*. The sub-commentary gives the explanation ‘ford’, but also alludes to the idea of ‘shining’. The matter is further complicated by the fact that the Pāli

<sup>10</sup> See W.B. Bollée, “Buddhists and Buddhism in the earlier literature of the Śvetāmbara Jains”, in L. Cousins et al., eds., *Buddhist Studies in Honor of I.B. Horner*, Dordrecht 1974, 27-39 (p. 36, n. 2).

<sup>11</sup> *kun-tu-bdag-po*. See *Ratnāvalī*, I, 93-95. Cf. *Yuktiṣaṣṭika*, 34.

<sup>12</sup> T, XXXII, 495b, 1.15.

<sup>13</sup> T, I, 102c, 1.17.

<sup>14</sup> T, I, 548b, 1.11.

<sup>15</sup> I, 223, 12.

<sup>16</sup> I, 329, 31.

<sup>17</sup> The Pali Text Society editions of the *Dīrgha-nikāya* and its commentary both read *-paha*, but this seems to be due to the confusion of *ha* and *bha* in the Sinhalese script.

grammarians Aggavaṃsa refers in one place in his grammar to *sabbato-pabha* as an epithet of *nibbāna*, but in another place refers to *sabbato-papha* as an example of the sound change *-p- > -ph-*.<sup>18</sup> It would seem likely that this is the sound change to which Buddhaghosa was referring, and we must therefore assume that there has been an error in the manuscript translation for this word - an assumption which is borne out when we note the variation in the readings in the various editions.

The most probable explanation for all this confusion is that in a version earlier than any of those available to us today the epithet had the form *sabbato-paha* or *sabbato-pahu*, i.e. it was composed in, or had been transmitted through, a dialect where aspirated stops developed to *-h-*, and where the nominative singular of short *-a* stems could be in *-o* or *-u*. Those translating into Pāli or Sanskrit were, therefore, faced with the problem of deciding how to present the word in their own language or dialect, and how to explain it. The Pāli tradition came up with three solutions: to change *-paha* to *-papha*, to change *-paha* to *-pabha*, or to change *-pahu* to *-pabhu*. The first was explained as *-papa* ‘ford’, with the change of *-p- > -pha-*; the second as *-pabha* ‘shining’, and the third as *-pabhūta* ‘abundant, having power’. Not all of these are attested in the canonical texts as we have them, but the commentarial traditions retained them in their exegesis.

Other traditions, at least those which are available to us now, seem not to have approved of, or perhaps thought of, the idea of *-papha* ‘ford’. It is not clear what the reading was in the Sanskrit or (Prakritized Sanskrit) versions underlying the Chinese *āgamas*. The version available to the *Dīrgha-āgama* redactor was clearly capable of interpretation as *-prabha*, which accounts for the translation found there. It seems likely that the version available to the *Madhyama-āgama* redactor was not capable of such an interpretation or it would surely have been translated in the same way as in the *Dīrgha-āgama*. Whatever it was, it seems to have been beyond the redactor’s ability to translate, which probably accounts for his omitting it. The version available to Nāgārjuna either contained the word *-pabhu*, or was capable of being so interpreted, which accounts for his including this form of the epithet in his *Ratnāvalī*.

I am not a scholar of Buddhism, and I must confess that I do not have any great interest in the subject, and know little or nothing about it. I would, however, describe myself as a scholar of Pāli, even if I discover each year that I know less and less about the subject, and increasingly find that I accept less and less of whatever I thought I understood years ago. I regard my part in the connection between Pāli philology and Buddhist studies as being that of a consultant, and over the years I have had an extensive correspondence with those who wish to know whether the suggestions or proposals which they wish to make about Buddhism, based upon Pāli sources, are tenable and viable. To such inquiries I have occasionally had to say that, relying on the knowledge which I have of the subject, their suggestions are impossible or, rather, very unlikely (it is hard to be certain that anything is impossible in the field of Middle Indo-Aryan studies). Sometimes I can emphatically support the suggestion, and even give additional evidence. Most of the time, however, I can say little more than “Maybe”, which is sufficient for them, they believe, to go ahead.

---

<sup>18</sup> *Saddanīti* 70, 20 and 622, 21.

To return to the point which I tried to make at the beginning of this paper, it must be said that the Pali Text Society is well aware of the deficiencies of many of its editions, and, inevitably, of the translations based upon them. The problem is to know what to do about it. Faulty editions do not correct themselves by mere wishful thinking, and there is a desperate shortage of those who are both qualified to make satisfactory editions of Pāli texts and also willing to correct earlier editor's work rather than make an edition of some newly discovered work which they hope will have an earthshaking effect upon the world of Pāli and Buddhist studies when it appears. Quite often the amount of correction required in the old editions is so great that a new edition, rather than a corrected edition is required. When money is short, or workers lacking, then the Pali Text Society's general editor has himself, on occasion, made all the corrections that can be done by adding or removing diacritical marks or punctuation marks, with ink and whitener respectively. I have personally spent many hours in this way, preparing works for printing. When the Society decided to print the texts so arbitrarily deprived of hyphens by the Founder's widow, it fell to my lot to put them all back in by hand, since it would cost a large sum of money, inevitably reflected in the selling price of the book, if the printer had done it. Sometimes one's plans are upset by well-meaning people. I once spent many hours correcting a copy of a particular work for reprinting, only to find when I received a copy of the reprint that an over-zealous sub-editor, appalled at the number of handwritten corrections in the copy sent him for photographing, had searched high and low to get a 'clean', i.e. uncorrected, copy which he proceeded to send to the printers in place of the copy upon which I had worked so hard.

If the situation is to be improved, then action must be taken to increase the number of philologists working in the field of Buddhist studies. It is perhaps going too far to say that there is no shortage of those wishing to work in the field of Chinese and Tibetan Buddhism, but certainly there are recruits coming forward in those fields. As I have, however, tried to make clear, Chinese and Tibetan, by themselves, are not sufficient for those who wish to specialize in Hīnayāna Buddhism. Sanskrit is clearly essential, and so too is competence in Middle Indo-Aryan, by which I mean not just Pāli but the whole range of the dialects coming under that heading, including those used by the Jains for their canonical and commentarial texts. As I have emphasized, the texts which we have in Pāli, no less than the Hīnayāna canonical texts which we have in Sanskrit, are translations from other dialects of Middle Indo-Aryan, and to understand how Pāli and Sanskrit texts came to be in the form of which we have them today we have to know as much as we can about those other dialects. Unless we can attract recruits to the field of Middle Indo-Aryan studies, then the supply of those with the necessary knowledge will dry up, and articles and books about Buddhism will continue to be written by those who cannot handle the language themselves and will consequently, of necessity, be dependant upon the unsatisfactory texts and translations of which, with a few notable exceptions, we have at the moment.